

**CLAIMS**What is claimed is:

1. A method comprising:  
receiving data from a measurement on a biological sample;  
cutting data at, at least one time interval ;  
aligning data based on the cutting to form an  $n$ -dimensional data set;  
and  
storing an  $n$ -dimensional sub-region of data elements from the  $n$ -dimensional data set, wherein a size of a dimension of the  $n$ -dimensional sub-region is less than a size of the dimension of the  $n$ -dimensional data set.
2. The method of claim 1, wherein the  $n$ -dimensional data set exceeds an addressable memory limit for a data processing system.
3. A method comprising:  
dividing an  $n$ -dimensional data set into  $n$ -dimensional sub-regions, wherein a size of a dimension of an  $n$ -dimensional sub-region is less than a size of the dimension of the  $n$ -dimensional data set, the  $n$ -dimensional data set is obtained from a measurement made on a biological sample; and  
storing the  $n$ -dimensional sub-regions on a computer readable medium, the computer readable medium is accessible to a data processing system.
4. The method of claim 3, wherein the  $n$ -dimensional data set exceeds an addressable memory limit for the data processing system.

5. The method of claim 3, wherein a size of the  $n$ -dimensional sub-region is set equal to a size of one page of memory.
6. The method of claim 3, wherein a size of each dimension of the  $n$ -dimensional sub-region is a power of two.
7. The method of claim 6, further comprising;  
choosing the size of each dimension to balance uniformity in all dimensions against wasted storage space.
8. The method of claim 7, wherein a size of a dimension that will be traversed most rapidly during an analysis is increased relative to the size of the other dimensions.
9. The method of claim 6, wherein the size of each dimension is chosen to avoid excessive wasted memory while also minimizing page swapping.
10. The method of claim 9, wherein the largest dimension is chosen to be the dimension that is traversed most rapidly during an analysis.
11. The method of claim 3, wherein a detection is used to measure the property and the detection is selected from the group consisting of flame ionization detection (FID), thermal conductivity detection (TCD), electron capture detection (ECD), flame photometric detection (FPD), hall electrolytic

conductivity detection, laser-induced fluorescence (LIF) detection, ultraviolet (UV) transmission detection, a transmission detection, autoradiological imaging detection, visible wavelength reflectivity imaging with a stain detection, visible wavelength reflectivity imaging without a stain detection, non-visible wavelength reflectivity imaging with a stain detection, and non-visible wavelength reflectivity imaging without a stain detection.

12. The method of claim 3, wherein the measurement is obtained with an analytical unit.

13. The method of claim 12, wherein the analytical unit is selected from the group consisting of liquid chromatography (LC), gas chromatography (GC), capillary electrophoresis (CE), solid phase extraction, gel chromatography (gelC), open-bed chromatography (planar chromatography), high performance liquid chromatography (HPLC), absorption chromatography, ion-exchange chromatography, normal phase chromatography, reverse phase chromatography, size exclusion chromatography, capillary electrophoresis (CE), capillary zone electrophoresis (CZE), capillary gel electrophoresis (CGE), capillary isoelectric focusing (CIEF), isotachopheresis (ITP), electrokinetic chromatography (EKC), micellar electrokinetic capillary chromatography (MECC OR MEKC), capillary electrochromatography (CEC), non-aqueous capillary electrophoresis (NACE), gel chromatography (GC), a one-dimensional gel method, a two-dimensional gel method, a device acting as a method of gel chromatography, thin layer chromatography (TLC), paper chromatography, affinity chromatography, mass spectrometer, time-of-flight

(TOF) mass spectrometer, magnetic sector mass spectrometer, quadrupole mass spectrometer, ion trap mass spectrometer, ion cyclotron resonance mass spectrometer, Fourier transform ion cyclotron resonance (FTICR) mass spectrometer, mass spectrometer with electrospray ionization (ESI), matrix-assisted laser desorption/ionization (MALDI) mass spectrometer, surface enhanced laser desorption/ionization (SELDI) mass spectrometer, charge induced dissociation (CID) mass spectrometer, and in source decay mass spectrometer.

14. The method of claim 3, wherein the measurement is obtained by combining a plurality of analytical units.

15. The method of claim 14, wherein at least one of the plurality of analytical units is a capillary electrophoresis process.

16. The method of claim 14, wherein at least one of the plurality of analytical units is a means for performing capillary electrophoresis.

17. The method of claim 14, wherein at least one of the plurality of analytical units is a chromatography process.

18. The method of claim 14, wherein at least one of the plurality of analytical units is a means for performing chromatography.

19. The method of claim 14, wherein at least one of the plurality of analytical units is a mass spectroscopy process.
20. The method of claim 14, wherein at least one of the plurality of analytical units is a means for performing mass spectroscopy.
21. The method of claim 3, further comprising:  
swapping a second  $n$ -dimensional sub-region with a first  $n$ -dimensional sub-region; and  
utilizing the second  $n$ -dimensional sub-region in a mathematical operation.
22. The method of claim 21, further comprising;  
traversing through the  $n$ -dimensional data set in sub-region order.
23. The method of claim 22, wherein the traversing is used in a mathematical operation that performs pattern recognition on the  $n$ -dimensional data set.
24. The method of claim 21, wherein the mathematical operation is used to identify a protein associated with the biological sample.
25. The method of claim 21, wherein the mathematical operation is used during an analysis of a biological sample and a process used during the

analysis is selected from the group consisting of aligning, re-sampling, averaging, noise suppression, de-convolution, and peak-finding.

26. The method of claim 3, further comprising:

using metadata to search for a data element from the  $n$ -dimensional sub-regions, wherein the metadata enables the search to be performed on a subset of the  $n$ -dimensional sub-regions.

27. The method of claim 26, further comprising:

retrieving the data element based on the using.

28. A computer readable medium having stored thereon a data structure comprising:

a first field containing data representing a data value of an  $n$ -dimensional array, wherein the data value corresponds to a measurement made on a biological sample;

a second field containing data representing an  $n$ -dimensional sub-region number, the data value is assigned to the  $n$ -dimensional sub-region number; and

a third field containing data representing an offset into the  $n$ -dimensional sub-region number that corresponds to a location of the data value.

29. The computer readable medium of claim 28, further comprising;

a fourth field containing meta-data.

30. The computer readable medium of claim 29, wherein the meta-data provides information on a property of the *n*-dimensional sub-region.

31. The computer readable medium of claim 30, wherein the property is selected from the group consisting of a boundary, a maximum data value, a minimum data value, and a peak in the data.

32. A method comprising:

reading an *n*-dimensional data set, the *n*-dimensional data set represents a measurement of a property of a biological sample and the *n*-dimensional data set exceeds an amount of addressable storage associated with a data processing system;

dividing the *n*-dimensional data set into *n*-dimensional sub-regions, where a size of a dimension of an *n*-dimensional sub-region is less than a size of the dimension of the *n*-dimensional data set;

storing the *n*-dimensional sub-regions; and

performing mathematical operations on the *n*-dimensional sub-regions, wherein pattern recognition is applied to the *n*-dimensional data set through the performing.

33. The method of claim 32, wherein the measurement is obtained by combining at least two analytical units.

34. The method of claim 32, wherein a size of the  $n$ -dimensional sub-region is set equal to a size of one page of memory.
35. The method of claim 32, wherein a size of each dimension of the  $n$ -dimensional sub-region is a power of two.
36. An apparatus comprising:
- an analytical unit, the analytical unit is configured to make a measurement of a property of a biological sample, wherein an  $n$ -dimensional data set is obtained from the measurement;
  - a storage device; and
  - a processor programmed to:
    - divide the  $n$ -dimensional data set into a plurality of  $n$ -dimensional sub-regions, wherein a size of a dimension of an  $n$ -dimensional sub-region is less than a size of the dimension of the  $n$ -dimensional data set;
    - and
    - maintain in the storage device the plurality of  $n$ -dimensional sub-regions.
37. The apparatus of claim 36, further comprising:
- a prepended analytical unit, the prepended analytical unit is in communication with the analytical unit and the biological sample passes from the prepended analytical unit into the analytical unit, wherein the apparatus is configured to measure a property of the biological sample.



38. The apparatus of claim 37, wherein the property is selected from the group consisting of a mass spectrum, an electric current, and a general property.

39. The apparatus of claim 36, wherein a first dimension of the  $n$ -dimensional data set is obtained by cutting an output of the analytical unit at a time interval into segments and then aligning the segments of the output to form a first dimension of the  $n$ -dimensional data set.

40. The apparatus of claim 36, wherein the measurement is obtained by combining a plurality of analytical units.

41. The apparatus of claim 40, wherein at least two of the plurality of analytical units is combined in series.

42. The apparatus of claim 40, wherein at least two of the plurality of analytical units is combined in parallel.

43. The apparatus of claim 36, wherein the  $n$ -dimensional data set exceeds the processor's addressable memory limit and the processor is further programmed to:

perform mathematical operations on the  $n$ -dimensional data set by accessing data from the  $n$ -dimensional sub-regions.

44. The apparatus of claim 43, wherein a protein contained in the biological sample is identified.

45. The apparatus of claim 43, wherein pattern recognition is performed on the  $n$ -dimensional data set.

46. A computer readable medium containing executable computer program instructions, which when executed by a data processing system, cause the data processing system to perform a method comprising:

dividing an  $n$ -dimensional data set into  $n$ -dimensional sub-regions, wherein a size of a dimension of an  $n$ -dimensional sub-region is less than a size of the dimension of the  $n$ -dimensional data set, the  $n$ -dimensional data set is obtained from a measurement made on a biological sample; and

storing the  $n$ -dimensional sub-regions on a computer readable medium, the computer readable medium accessible to a data processing system.

47. The computer readable medium, as set forth in claim 46, wherein the  $n$ -dimensional data set exceeds an addressable memory limit for the data processing system.

48. The computer readable medium, as set forth in claim 46, the method further comprising:

swapping a second  $n$ -dimensional sub-region with a first  $n$ -dimensional sub-region; and

utilizing the second  $n$ -dimensional sub-region in a mathematical operation.

49. The computer readable medium, as set forth in claim 48, the method further comprising:

traversing through the  $n$ -dimensional data set in sub-region order.

50. The computer readable medium, as set forth in claim 48, wherein the mathematical operation is used to identify a protein associated with the biological sample.

51. An apparatus comprising:

means for storing an  $n$ -dimensional data set, wherein the  $n$ -dimensional data set represents a measurement of a property of a biological sample;

means for performing mathematical operations on the  $n$ -dimensional data set when the  $n$ -dimensional data set exceeds an addressable memory limit of a data processing system.

52. The apparatus of claim 51, further comprising:

means for measuring the property of the biological sample.

53. The apparatus of claim 51, wherein the mathematical operations are used to apply pattern recognition to the biological sample.